# Ordinally Efficient and Reliable Social Choice: The Pluri-Borda Rule

Klaus Nehring*

February 28, 2018

## Abstract

Ordinally efficient choice correspondences are characterized as maximing a weighted average over comparison alternatives of the number of agents preferrring an alternative over any comparison alternative. The Borda rule is the classical example, assign equal weight to all comparison alternatives. However, the Borda rule suffers from severe problems of unreliability, as it is highly and implausibly sensitive to the inclusion of Pareto inferior and of minor variants ("clones") in the feasible set.

Reliability can be ensured, however, by an appropriate choice of weights that depends on the profile of preferences. This is achieved, for example, by weighting alternatives in proportion their "plurality', i.e. frequency as top choices, and induces the "Plurality-weighted Borda rule", or "Pluri-Borda rule" for short. While this weighting rule is not claimed to the ideal one, it has many attractive properties and comes with a transparent axiomatic characterization.

We summarize the analysis of the paper with an Arrowian possibility result, according to which the Pluri-Borda rule is able to jointly satisfy ordinal efficiency and reliability axioms on social choice, unlike any (familiar) rule in the literature.

*University of California, Davis

# 1 Introduction

We will adress the following version of the Arrowian social choice question. Given a set of feasible alternatives (agenda) and a profile of ordinal preference rankings over the agenda, which choices are socially best if all individuals count equally? Preferences are assumed to be known here, the agenda given. The Arrowian question asks about what the social optimum is, not about how it can be achieved. In particular, issues of strategic behavior are set aside as a question of mechanism design. The Arrowian question aims to determine what maximizes "aggregate ordinal welfare", or, in slightly different words, "aggregate preference satisfaction". Ex hypothesis, ordinal welfare must be determined from the preference relations of the individuals themselves, without recourse to any kind of additional information about the pyschological state of individuals that would permit interpersonal welfare comparisons of some kind.

We submit that a satisfactory answer to the Arrowian question has to address two issues in paricular. The proper content of "ordinalism', and a proper acknowlegment and treatment of the concerns with the potential unreliability of the output of a welfare criterion. Broadly speaking , reliability issues arise since full choice consistency is not achievable to the existence of majority cycles and related phenomena. In particular, a choice correspondence may specify the choice of one alternative, say $a$, at a particular profile, but, upon removal of the unchosen alternative $c$, may flip its choice to $b$. In principle, there is nothing particularly mysterious or questionable about such flips, since the inclusion of $c$ changes the information available for the evaluation of the non$-c$ elements, and so *may* reasonably change that evaluation. But it need not: the information obtainable via $c$ may have no, or only little, or only tenous bearing on that evaluation. The most prominent example is that of $c$ being a "clone" (intuively: 'minor variant') of some other alternative $c'$ already included in the agenda. Following Tideman (1987), a significant literature requires that the inclusion of additional clones should not be allowed to change the social choice. Another example, which will play a more central role in this paper, is the inclusion of alternatives that are Pareto dominated by alternatives in the original agenda. Positional scoring rules, and the Borda rule in particular, are highly susceptible to the inclusion of such apparently irrelevant alternatives, and would thus count as highly unreliable.

**Outline of Argument** While one would ultimately want to resolve the Arrowian question by selecting a unique alternative (up to non-generic ties), it is helpful to initially aim lower and asking whether, on the basis of a particular profile at a given agenda, one can identify pairs of alternatives that can be compared on 'straightforward' ordinalist grounds (beyond the Pareto criterion) and thereby at least identify some alternatives that are "clearly inadmissible".

We follow Dutta-Laslier (1999) in adopting *ordinal dominance* as such a baseline comparison: Alternative $a$ ordinally dominates alternative $b$ if $a$ is preferred to $b$ by a majority, and if, in the comparison to any third alternative, $a$ is preferred to $c$ by strictly more individuals than $b$ is to $c$. Alternatives not ordinally dominated by some other feasible alternative are called *"ordinally admissible"*. It is easy to see that Condorcet winners (if they exist) and Borda winners are ordinally admissible, while plurality winners need not be.[1]

Additional mileage is gained by extending this requirement to the absence of domination by randomized acts, refining the notion of ordinal admissibility to that of *"ordinal efficiency"* of alternatives and correspondences. A basic characterization result shows that a choice correspondence is ordinally efficient if, for every aggregation problem $(A, \mu)$, there exists a weight vector over alternatives such that the chosen alternatives maximize the weighted average of majority margins.

In view of this result, one can turn the problem of determining the best ordinally efficient alternative(s) into a problem of determining the appropriate weights as a function of the aggregation problem $(A, \mu)$. We will refer to such weighting functions as "indices" $\rho$, and the induced choice correspondence as "generalized Borda rules" $B_\rho$. The simplest example of such a rule is, evidently, the classical Borda rule whose index assigns equal weight to all alternatives. Notwithstanding its impressive axiomatic credentials, the Borda rule faces serious issues concerning the reliability of its out. As we illustrate by example, any Pareto efficient alternativ can become a Borda winner by enriching the agenda with a sufficient number of appropriately chosen Pareto dominated alternatives. It seems extremely implausible that, in the enriched agenda, the Borda rule determines the correct ordinal welfare optimum. Indeed, it is hard to see that anything material has changed at all with respect to the evaluation of the Pareto efficient alternatives. We thus conclude that the Borda rule is prone to serious error, especially if the number of alternatives is moderate or large.

But within the framework of GBRs introduced here, it is quite straightforwar do address such issues of reliability by imposing appropriate reliability conditions on the index itself. For example, the potentially distortionary role of Pareto dominated alternatives can be neutralized simply by giving such alternatives an index weight zero. Likewise, potential distortions that might arise from the doublecounting of "clones" might be neutralized by distributing weights appropriately over entire clusters of clones.

In the third part of the paper, we propose one particular GBR associated with a particular simple and natural weighting function, the plurality index. The plurality index assigns weights to alternatives in proportion to the number of individuals for whom this alternative is the top choice. We provide a simple axiomatization. The GBR derived from the plurality index will be referred as the "Plurality-Based Borda" rule, or "Pluri-Borda" rule for short. While we do not claim the Pluri-Borda rule to *the* "ideal" (normatively correct) ordinal welfare criterion point out to some potential shortcomings, the Pluri-Borda rule does seem to do a very credible job in most aggregation problems. Its simplicity

[1] Dutta-Laslier (1996) consider and discuss stronger versions of this relation.

and tractability are an additional, non-minor bonus.

**Plan of the Paper** Framework and Notation are introduced in section 2. Section 3 defines the central notions defining our approach to "ordinal welfarism". Its main result is a characterization of ordinally efficient correspondences as subcorrespondes of generalized Borda rules. Generalized Borda rules are defined in terms of the matrix of majority margins. In section 3.3, we show how this translates into a representation that is additively separable in individual generalized Borda scores. Section 4 illustrates the severe reliability deficit of the classical Borda rule by means of an example. Section 5 introduces a couple of regularity and reliability requirements on the problem-dependent weighting function (index), and derives some general choice implications. Section 6 then takes a look at the special but important domain of single-peaked preferences on a line, demonstrating that any "regular" Generalized Borda rule will at some profiles not choose the Condorcet winner. Section 7 defines the Plurality index and the associated Pluri-Borda rule, and provides a simple yet illuminating axiomatization. The argument is summarized in Section 8 which presents an "Arrowian possibility result" which shows that the Pluri-Borda is the first rule in the literature to jointly satisfy a number of choice axioms that, we submit, are to be satisfied by any choice correspondence that purports to reliably pin down a properly ordinal welfare criterion. In section 9, we round off the argument by asking whether Generalized Borda rules are consistent with the Condorcetian premise that Condorcet winners are to be chosen whenever they exist. Indeed, we show that the Essential Set introduced by Dutta-Laslier (1999) and axiomatized by Laslier (2000) is such an example.[2] It has excellent reliability properties. However, in contrast to the Pluri-Borda rule, it fails to be "resolute" as it selects at least three alternatives in the absence of Condorcet winner.

[2]The Essential Set is in fact the only GBR (other than the Borda rule) that has been defined in the prior literature. The weights supporting the essential set are the "maximal lotteries" introduced by Kreweras (1963) and Fishburn (1984) and axiomatized by Brandl et al. (2016).

## 2 Framework and Notation

We will employ a variable-agenda, variable-population framework.

- Let $\mathfrak{A}$ be an abstract (finite or infinite) set interpreted as the 'universe' of possible alternatives.

- A *choice set* or "agenda" is a finite subset of $\mathfrak{A}$; the family of all possible agendas – all finite subsets of $\mathfrak{A}$ – is denoted by $\mathcal{F}(A)$.

- Each individual $i \in I$ has a preference relation described by a linear order $P_i$ on $A$. For any given $A \in \mathcal{F}(\mathfrak{A})$, let $\mathcal{L}(A)$ denote the set of linear orders on $A$.

- A *profile* $\mu$ is a rational-valued probability distribution on $\mathcal{L}(A)$, with $\mu_{P.}$ denoting the relative frequency of individuals with preference ordering $P$. (That is, if one starts from a profile in the usual manner as an $I-$tuple of preferences, $\mu_P = \frac{|\{i:P_i=P|}{|I|}$.) Thus our description of a profile builds in *anonymity.* By considering relative rather than absolute frequencies, we have slightly restricted the generality by imposing an additional homogenity restriction. This has been done for expositional purposes. (All concepts and results could be rephrased on the basis of profiles described by absolute frequencies).

- The set of profiles is given by the rational-valued probability simplex on $\mathcal{L}(A)$ denoted by $\Delta_{\mathbb{Q}}(\mathcal{L}(A))$. The support of a profile will be denoted by $\mathcal{P}(\mu)$; i.e. $\mathcal{P}(\mu) := \{P \in \mathcal{L}(A) : \mu_P > 0\}$.

- An *aggregation problem* is a pair $(A, \mu) \in (\mathcal{F}(\mathfrak{A}), \Delta_{\mathbb{Q}}(\mathcal{L}(A)))$; the collection of aggregation problems is denoted by $\mathfrak{D}$. It is sometimes convenient to describe aggregation problems by reference to distributions $\nu$ over rankings on a larger agenda $B$. Then the pair $(A, \nu)$ denotes the aggregation problem $(A, \mu) \in \mathfrak{D}$, where $\mu_P = \sum_{P':\mathrm{Rest}_{|A}(P')=P} \nu_{P'}$

- A social *choice correspondence* $C$ maps voting problems $(A, \mu) \in \mathfrak{D}$ to non-empty subsets of $A$.

- For any $P$, $top(P)$ denotes its maximal element. For any $(A, \mu) \in \mathfrak{D}$, $tops(A, \mu) := \{top(P) : \mu_P > 0\}$. Likewise, $\pi(A, \mu)(a) := \sum_{P \in \mathcal{L}(A)} \{\mu_P : top(P) = a\}$. So, the "plurality measure" or "plurality index" is the image measure of $\mu$ under *top*, $\pi(A, \mu) = \mu \circ top^{-1}$.

# 3 Generalized Borda Rules

## 3.1 Ordinal Dominance

The following discussion make essential use of the (common) notion of *majority margin* $M(a,b)$ between two alternatives.

$$M(a,b) := \mu(\{P : aPb\}) - \mu(\{P : bPa\}).$$

Note that, by construction, the "matrix" $M$ is skew-symmetric, i.e. $M(a,b) = -M(b,a)$ for all $a,b \in A$.

Without change in substance, we could have developed the material on the basis of the (gross) majorities $\overline{M}(a,b)$ of individuals favoring $a$ over $b$, with indifference (identity) splitting fifty-fifty.

$$\overline{M}(a,b) := \mu(\{P : aPb\}) \text{ if } a \neq b, \text{ and } \overline{M}(a,b) = \frac{1}{2} \text{ if } a = b.$$

It is easily checked that $\overline{M} = \frac{1}{2}(1+M)$, so that $M$ and $M'$ are positive affine transformations of each other, hence equivalent for all purposes relevant here.

Given an aggregation problem $(A,\mu)$, which alternatives are "better in the aggregate" on the basis of the ordinal preference comparisons only? With $|A| = 2$, the ordering by majority comparison is arguably compelling. It could be further backed up by May's theorem.

What about agendas with cardinality $|A| > 2$. A 'Naive' proposal would extend the direct majority comparison to non-binary agendas. But this cannot work due to the possibility of Condorcet cycles. One could put the error down to forced neglect of *indirect* ordinal preference information. So if $M(a,b) > 0$ while $M(b,c) > 0$ and $M(a,c) < 0$, the indirect comparison favors $b$ over $a$. Heuristically, if the indirect majority margins are strong enough and "deserve" sufficient weight, they should outweight the direct comparison.

A priori, it is not obvious how to settle the trade-off between the direct and the indirect comparisons in general, and this can be viewed as one of the main sources of the difficulites of voting theory. Here we begin by focusing on cases where there are no such trade-offs to make. Thus, say that *a ordinally dominates b* in the aggregation problem $(A,\mu)$ iff, for all $z \in A$,

$$M(a,z) > M(b,z).$$

The alternative $a$ is *ordinally admissible* in $(A,\mu)$ if there does not exist any other alternative $b \in A$ ordinally dominating it. Let their set be denoted by $OA(A,\mu)$. A social choice correspondence $C$ is *ordinally admissible* if $C(A,\mu) \subseteq OA(A,\mu)$ for all $(A,\mu) \in \mathfrak{D}$.[3]

[3] One might be drawn to want to refine ordinal admissibility by considering "weak" instead of "strict" dominance, but this raises a host of issues. For example, naturally defined and well-motivated ordinally admissible choice functions such as the Essential Set defined below in section 9 contain weakly dominated alternatives. Dutta-Laslier (1999) conclude their thorough analyis of the issue by suggesting that one may need to settle for this.

Note that, if $A = \{a, b\}$, then $a$ ordinally dominates $b$ in $\{a, b\}$ iff $M(a, b) > 0$. By contrast, for non-binary agendas , the partial ordering defined by the ordinal dominance relation is often quite weak, admitting a range of (non-equivalent) multiple admissible alternatives. However, in special, yet illuminating aggregation problems $(A, \mu)$, only a single alternative may be ordinally admissible.

**Example 1** *Let $A = \{a\} \cup B$, with $|B| \geq 3$. The profile $\mu$ is given by the following two conditions. First, the marginal distribution of $\mu$ on rankings over $B$ is the uniform distribution. Second, the alternative $a$ is ranked second by all $P \in \mathcal{P}(\mu)$. From the first condition, it is immediate that, for all $b, b' \in B$, $M(b', b) = 0$. By the second condition, $a$ is ranked above $b$ exactly by those individuals whose top is not $b$. These have measure $1 - \frac{1}{n}$, entailing a majority margin of $M(a, b) = 1 - \frac{2}{n} > 0$ for all $b \in B$. Thus $a$ ordinally dominates any $b \in B$.*

From this Example, it follows immediately that no *top-confined choice correspondence* is ordinally admissible, where "top-confined" means that the correspondence chooses only alternatives that are some individual's top. Examples are Plurality rule, Plurality with a run-off, and Single-Transfareable vote. This strengthens the common critique of these rules as not being Condorcet consistent.

Ordinal dominance is not really a new concept although it seems to be studied rarely; it is a weakening of the covering relation considered by, e.g. Dutta-Laslier 1999. So the ordinally admissible set is contained in the Pareto efficient set and contains the uncovered set.

How should one select among ordinally admissible alternatives? A classical proposal is Borda's which can be defined as maximizing the average majority margin over all comparison alternatives. The Borda Rule $C^{Borda}$ is given by

$$C^{Borda}(A, \mu) = \arg\max_{a \in A} \frac{1}{n} \sum_{e \in A} M(a, e).$$

The Borda Rule is a special case of a scoring rule. For a given cardinality of the agenda, a scoring function $t$ is a non-increasing mapping from the set $\{1, .., |A|\}$ to the unit interval $[0, 1]$. It is usually assumed without loss of generality that $t(1) = 1$ and $t(|A|) = 0$. The aggregate score of an alternative being defined as $t(a, \mu) = \sum_{P \in \mathcal{L}(A)} (t(\{|b : bPa|\} + 1)\, \mu(P)$, the associated scoring rule $C^{t-score}$ is given by is defined, for fixed agenda $A$,

$$C^{t-score}(\mu) = \arg\max_{a \in A} t(a, \mu).$$

**Proposition 2** *Let $A$ be any fixed agenda. A scoring rule $C^{t-score}$ is ordinally admissible if and only if it is the Borda rule.*

This result is reminiscent of results that identify the Borda rule as the unique scoring rule that is C2 – that is, depends only on the majority matrix $M$. But ordinally admissibility does not entail C2, nor is it entailed by it. It is also closely related to characterizations of the Borda rule involving Cancellation. So it is no surprise given the literature, but good to have here to flesh out the overall picture.

### 3.2 Ordinal Efficiency

One can naturally extend the ordinal dominance to randomized ("lotteries") $p \in \Delta(A)$ by considering expected majority margins[4]

$$M(p,q) = \sum_{a,b \in A} p_a M(a,b) q_b.$$

It is easy to check that skew-symmetry of $M$ on $A \times A$ entails skew-symmetry of $M$ on $\Delta(A) \times \Delta(A)$, i.e. $M(p,q) = -M(q,p)$ for all $p, q \in \Delta(A)$.

So, identifying the alternative $a$ with the degenerate lottery $\delta_a$, say that the lottery $p$ **ordinally dominates** $a$ iff (in $A$) iff, for all $z \in A$,

$$M(p,z) \geq M(a,z),$$

with at least one strict inequality. An alternative $a \in A$ is **ordinally efficient** if there does not exist a lottery $p \in \Delta(A)$ such that $p$ ordinally dominates $a$.

Ordinally efficient alternatives maximize a weighted average of majority margins. The following result follows from standard separation arguments for convex sets.

**Proposition 3** *The alternative $a \in A$ is ordinally efficient in $A$ if and only if there exists a weight vector $w \in \Delta(A)$ such that, for all $b \in A$,*

$$\sum_{e \in A} w_e M(a,e) \geq \sum_{e \in A} w_e M(b,e). \tag{1}$$

**Proof.** For any lottery $p \in \Delta(A)$, let $\mathbf{m}_p$ denote the vector $(M(p,e))_{e \in A} \in [-1,+1]^A$. Let $\mathcal{M}$ denote the closed convex set $\{\mathbf{m}_p\}_{p \in \Delta(A)}$. Evidently, $a$ is ordinally efficient in $A$ iff the open convex set $\{\mathbf{m} \in \mathbb{R}^A : \mathbf{m} > \mathbf{m}_a\}$ does not intersect $\mathcal{M}$. Thus, by a standard separation argument for convex sets, there exists $w \in \Delta(A)$ satisfying equation (1).

[4] Expected majority margins play a key role in the theory of "Maximal Lotteries" (Fishburn 1984, Brandl et al. 2016). In contrast to these accounts, we consider non-stochastic choice functions $F$. These can be canonically extended to (non-resolute) stochastic choice functions by requirement that a lottery be maximal at a profile if and only if all alternatives in its support are. That is, $p \in \overline{C}(A,\mu)$ iff $a \in C(A,\mu)$ for all $a \in$supp$(p)$.

The converse ("if") is straightforward. □

The following is an interesting illustration of the gain in normative strength achieved by the ordinal efficiency criterion.

**Corollary 4** *If $a$ is ordinally efficient, $a$ is not a Condorcet looser.*

**Proof.** By Proposition 3, there exists $w \in \Delta(A)$ satisfying equation (1) for the ordinally efficient $a$.

By the symmetry of the matrix of majority margins, $M(a,w) \geq M(w,w) = 0$. But, if $a$ is a Condorcet Loser, $M(a,b) < 0$ for all $b \neq a$. Hence $M(a,w) < 0$ unless $w = \delta_a$. In that case, $M(a,w) = 0 < M(b,w)$ for all $b \neq a$, contradicting equation (1). □

An interesting example of a choice rule that is (weakly) ordinally admissible but not (weakly) ordinally efficient is the minmax or Simpson-Kramer rule. It follows immediately that it is weakly (but not strictly) ordinally admissible. It is also well-known that it may select Condorcet losers. This is generally viewed as a serious defect of the minmax rule . The ordinal effiency criterion appears to be an attractive normative *argument* for excluding Condorcet losers.

### 3.3 Joint Ordinal Efficiency

Consider sets of alternatives $G$ as potential output of a choice correspondence at $(A,\mu)$. When can the set $G$ be regarded as potentially identifying the set of ordinally optimal choices? For this to be the case, not only needs each $a \in G$ be ordinally efficient individually, but the different $a \in G$ must be viewable as "jointly tied". One way to capture make this notion of "joint tie" is again via lotteries. Think of $G$ as a choice rule's "recommendation". That is, the rule permits choosing any in $G$ without pinning down which in a definite manner. In the execution of the rule, the implementor may choose any, but must finally choose one. This choice is, qua content of the recommendation, essentially arbitrary. So it could be made contingent on some external factor, such as day of the week; indeed, one can argue that some external factor must have come in to cause the choice of one over another. Such contingency introduces potential uncertainy. It is most directly captured by considering the choice of "lotteries" among the recommended alternatives.

If $G$ represents a genuine tie, any such lottery should be potentially optimal as well; in particular, it needs to be ordinally efficient. So we say refer to a set $G$ as **jointly ordinally efficient** for $(A,\mu)$ if, for no lottery $p \in \Delta(A)$ with support contained in $G$, there exist another lottery $q \in \Delta(A)$ such that $q$ ordinally dominates $p$. An choice correspondence $C$ is **ordinally efficient** if $C(A,\mu)$ is joint ordinally efficient for all $(A,\mu) \in \mathfrak{D}$.

From these definitions, it seems "obvious" that the joint ordinal efficiency of sets may be a stricter requirement than the ordinal efficiency of each element. It is indeed so, but not entirely trivial to verify, so we have chosen to record it as a formal result.

**Proposition 5** *There exist aggregation problems $(A, \mu)$ for which the set of ordinally efficient alternatives $OE(A, \mu)$ is not jointly efficient.*

Extending the argument of Proposition 3 above,

**Proposition 6** *The choice correspondence $C$ is ordinally efficient if and only if, for all $(A, \mu)$ there exists a weight vector $w^{(A,\mu)} \in \Delta(A)$ such that, for all $a \in C(A, \mu)$ and all $b \in A$,*

$$\sum_{e \in A} w_e^{(A,\mu)} M(a, e) \geq \sum_{e \in A} w_e^{(A,\mu)} M(b, e). \tag{2}$$

Note that the "joint optimality" of the sets $C(A, \mu)$ is reflected in maximality of the average majority margin with respect to a common weight vector $w^{(A,\mu)}$.

**Proof.** The implication from (2) to (1) is straightforward. To see the converse, take any agenda $A$ and any profile $\mu \in \Delta(A)$. Let $p$ be any lottery whose support equals $C(A, \mu)$. By Proposition 3 (applied without change to ordinally efficient lotteries rather than just alternatives), there exist a vector $w = w_{a,\mu} \in \Delta(A)$ such that,

$$M(p, w) \geq M(q, w) \text{ for all } q \in \Delta(A).$$

By linearity of $M$ in the first argument and the support assumption on $p$,

$$M(p, w) = M(a, w) \text{ for all } a \in C(A, \mu).$$

Combining these two (in)equalities, we infer that, for all $a \in C(A, \mu)$ and $b \in A$,

$$M(a, w) \geq M(a, w),$$

as asserted. □

Proposition 6 can be used in at least two different ways. On the one hand, one can use it as providing a technical criterion for *checking* whether a given choice correspondence is jointly ordinally efficient; the weight vectors $w^{(A,\mu)}$ then serve merely as 'witness' of this property, without interpretable meaning of their own.

On the other hand, a more farreaching reading, Proposition 6 to *determine* which j.o.e. choice correspondences "appropriately", or even best serve as criteria of "ordinal welfarism"; on this reading, ordinal welfarism boils down to an exercise in multi-criterion decision making, where each alternative is evaluated by the same set of $|A|$ criteria, where criterion $e \in A$ measures the size of the majority margin of $a$ versus $e$. The trade-off among these criteria is determined by the weight vector $w^{(A,\mu)}$. One can think of the weighted average $\sum_{e \in A} w_e^{(A,\mu)} M(a,e)$ of representing an "index" of aggregate welfare, with the $e \in A$ representing the different index components. Different indices will be characterized by different component weights. The Borda rule represents the simplest instance of it, with component weights equal at any profile.

To highlight the independent, determinative role of the weights, an **index weighting function** $\rho$ will be any function that assigns to any aggregation problem $(A,\mu)$ a non-negative vector of weights $\rho(A,\mu) \in \Delta(A)$. Any i.w. function $\rho$ induces a choice correspondence $B^\rho$ given by

$$B_\rho(A,\mu) := \arg\max_{a \in A} \sum M(a,e)\rho_e(A,\mu).$$

$B^\rho$ will be called the **generalized Borda rule** (GBR) based on index $\rho$.

To reflect the change in perspective, we can now restate Proposition 6 as follows.

**Theorem 7** 1. *A choice correspondence $C$ is jointly ordinally efficient if and only if $C \subseteq B_\rho$ for some index weighting function $\rho$.*

2. *A choice correspondence $C$ is a generalized Borda rule if and only if it is jointly ordinally efficient and inclusion-maximal among all choice correspondences with this property.*

We do not insist on inclusion-maximality as normatively mandated here. That, indeed, would entail that there was a perfect tie among all generalized Borda maximizers; but this need not be the case, as one might be able to invoke additional considerations to refine the selection.[5]

Formally, index weighting functions $\rho$ are the same as probabilistic social choice functions, as in Brandl et al. (2016). Indeed, from this perspective one can think of the GBRs as selecting alternatives with the highest expected majority margin $M(a, \rho(A,\mu))$ of $a$ over some "reference lottery" $\rho(A,\mu)$. While this adds an additional heurist angle on what a GBR does, it does not appear to be helfpul to determine what $\rho$ should be in the first place.

[5] We intend to adress the refinement issue further in future versions of the paper.

### 3.4 Scoring Characterization of Generalized Borda Rules

To connect the notion of GBRs to the usual, rank-based, understanding of Borda rules, it is helpful to provide a representation that is additively separable in individuals.

To do so, it is helpful to view GBRs as maximizing average gross majorities $\overline{M}(a, \rho(A, \mu))$. Recalling that $\overline{M} = \frac{1}{2}(M+1)$, this comes out to the same thing.

Abreviating $\rho(A, \mu)$ to $\rho$, evidently

$$\begin{aligned}\overline{M}(a, \rho) &= \sum_{e \in A} \mu(\{P : aPe\}) \rho_e + \frac{1}{2}\rho_a \\ &= \frac{1}{n}\sum_{i \in I}\left(\sum_{e \in A : aP_i e} \rho_e + \frac{1}{2}\rho_a\right).\end{aligned}$$

Thus, if we define the generalized Borda score $s_i(a; \rho(A, \mu))$ as

$$s_i(a; \rho(A, \mu)) := \left(\sum_{e \in A : aP_i e} \rho_e + \frac{1}{2}\rho_a\right),$$

one has

$$B_\rho = \arg\max_{a \in A} \frac{1}{n}\sum_{i \in I} s_i(a; \rho(A, \mu)).$$

Generalized Borda scores yield an intuitive imputed strength of preference (for $aP_i b$) of

$$s_i(a; \rho) - s_i(b; \rho) = \frac{1}{2}\rho(a) + \sum_{e : aP_i eP_i b} \rho(e) + \frac{1}{2}\rho(b).$$

Thus, an individual exhibits a strong preference for $a$ over $b$ to the extent that $a$ is preferred to alternatives $e$ of high total measure that are in turn preferred to $b$.

Note also that the range of scores is typically less than the unit interval, and will vary across individuals, even in size. [6]

[6] Indeed, it is given by $\left[\frac{1}{2}\rho(bottom(P_i)), 1 - \frac{1}{2}\rho(top(P_i))\right]$.

# 4 Whats is wrong with the Borda rule?

**Example 8** *We compare the choice in two aggregation problems* $(A, \mu)$ *and* $(A', \mu')$. *A consists of two alternatives only,* $A = \{a, b\}$, *with* $, 0 < \mu(ab) < \mu(ba)$. *Think of a and b as two "projects". Agents have the same preferences over A in both problems; that is, in the notation introduced in section 2,* $(A, \mu') = (A, \mu)$.

*The agenda* $A'$ *is obtained by adding to A a number of "slightly inferior variants"* $A''$ *of project a. Specifically, all individuals prefer a to any variant* $a' \in A''$, *and those individuals who prefer a to b prefer any variant* $a' \in A''$ *to b as well.*

Since the variants in $A'$ are all Pareto inferior to $a$, in $(A', \mu')$ the social choice boils down to one among $a$ and $b$, just as in $(A, \mu)$. Is there any reason that they should differ? More specifically, is there sufficiently strong reason to overturn the majority favoring $b$ over $a$ (in both problems) in the direction of $a$?

In principle, one might distinguish at least two informational scenarios. In scenario one, it is evident from the desription of the alternatives that they are "slightly inferior variants". For example, the $a'$ might result from $a$ by throwing away a few pennies here or there. In that case, the profile $\mu'$ in $A'$ could have been inferred (at least with high likelihood) from that in $(A, \mu)$, and it stands to reason that only additional feature of $(A', \mu')$ is the mere fact that the inferior variants have been included in the agenda, for whatever reason. In this scenario, there seems to be no argument at all for a change in choice towards $a$.

In a second scenario, the alternative $a' \in A''$ and $a$ are descriptively unrelated; for example, the alternatives might be identified by mere labels. In that scenario, one might argue that the fact $a$ performs better than $b$ in the indirect comparison with any $a' \in A$ is some, perhaps slight reason to upgrade the evaluation of $a$ in $A'$. How much is the case for switching the social choice to $a$ strengthen if $|A''|$ is large? This may be hard to assess 'intuitively', but the Borda rule gives a clearcut answer. Any majority of $b$ over $a$, however large, may be overturned by the introduction of a sufficient number of inferior variants. Specifically, a simple computation reveals that $C^{Borda}(A', \mu') = \{a\}$ whenever $|A''| + 1 \geq \frac{1}{\mu(ab)}$. Indeed this follows from observing the difference in Borda scores $s_i^{Borda}(A', \mu')(a) - s_i^{Borda}(A', \mu')(b)$ equals $1 - \frac{1}{|A''|+1}$ for those individuals with top $a$, while this difference is just $-\frac{1}{|A''|+2}$. Thus if $|A''| + 1 \geq \frac{1}{\mu(ab)}$, the aggregate Borda score

$$
\begin{aligned}
& s^{Borda}(A', \mu')(a) - s^{Borda}(A', \mu')(b) \\
= \; & \left(1 - \frac{1}{|A''| + 1}\right) \mu(ab) + \left(-\frac{1}{|A''| + 2}\right) \mu(ba) \\
> \; & \mu(ab) - \frac{1}{|A''| + 1} \left(\mu(ab) + \mu(ba)\right) = \mu(ab) - \frac{1}{|A''| + 1} \geq 0.
\end{aligned}
$$

In the framework of GBRs, there is a straightforward remedy for this apparent anomaly. Simply render all Pareto dominated alternatives irrelevant by setting their index weights to zero. Then, in the larger problem $(A, \mu')$, the index will remain concentrated on $\{a, b\}$, and the majority preference of $b$ over $a$ will be reinstated.

The Example could also be looked at in terms of Independence of Clones, since set $A'' \cup \{a\}$ is a cluster of clones at the preference profile $\mu'$; see the following section for formal definitions. Again, by adapting the index weights to reflect the structure of clones, one can insure reliability of the induced social choice in the presence of clones. In the example, the reasoning from Pareto Dominance and the reasoning from cloning would independently arrive at the same conclusion.

# 5 Minimal Index Requirements and their Choice Implications

In view of the general context-dependency of weights, it is a wide open question what appropriate or even "optimal" index weights would look like. We thus introduce a few basic requirements that indices should have intuitively; the list is not meant to be exhaustive, simply sufficiently specific to provide enough structure to allows to exhibit some common signature among this wide class of indices.

A good index should balance reliability with informativeness. To be informative, it should be sufficiently broad in its coverage. Intuitively, the index should include alternatives that have identifiable distinct merit. It thus plausible to require inclusion at least of those alternatives that are some individual's top choice.

**Axiom 9** ***(Dispersion)*** *For all* $(A, \mu) \in \mathfrak{D}$, $\rho(A, \mu)(a) > 0$ *whenever* $\pi(A, \mu)(a) > 0$.

We shall refer to the support of $\rho$ as the set of "relevant" alternatives, denoted by $rel(A, \mu)$, and to alternatives in the complement as "irrelevant" or "null".

If an alternative is null, its presence or absence should arguably be immaterial, as expressed by the following axiom.

**Axiom 10** ***(Null Consistency)*** *For all* $(A, \mu) \in \mathfrak{D}$, *if* $\rho(A, \mu)(a) = 0$, *then* $\rho(A \backslash \{a\}, \mu) = \rho(A, \mu)$.[7]

[7] Since the support of $\rho(A, \mu)$ is contained in $A \backslash \{a\}$ by assumption, $\rho(A, \mu)$ can be viewed as an element of $\Delta(A)$.

Since alternatives in their role as comparators are characterized here by the profile of individuals' preferences, presumably the index should depend continuously on the profile.

**Axiom 11** ***(Continuity)*** *For all $A$, $\rho(A, \mu)$ is continuous in $\mu$.*

These conditions are intended as "conditions of regularity" without much "edge". Note, in particular, that they are all satisfied trivially by the Borda index.

"Conditions of regularity" are to be contrasted with "conditions of reliability". For example, we have argued in the previous section that Pareto dominated are dispensable and potentially highly unreliable as comparators. Hence they should not be given any weight.

**Axiom 12** ***(Pareto Irrelevance)*** *For all $(A, \mu) \in \mathfrak{D}$, $\rho(A, \mu)(a) = 0$ whenever $a$ is Pareto dominated by some $b \in A$.*

Likewise, the index includes should avoid double-counting of very similar comparators by appropriate discounting of their weight. To ensure this, we shall make use of the notion of a "cluster of clones" common in the literature on voting and probabilistic social choice following Tidemann (1987). A set $B \subseteq A$ is a **cluster of clones** at $\mu$ if, for all $b, b' \in B$ and $a \in A \backslash B : aRb$ iff $aRb'$**.**

**Axiom 13** ***(Cloning Invariance)*** *If $B \subseteq A$ is a cluster of clones at $\mu$, then, for any $\varnothing \neq B' \subseteq B$, $\rho(B' \cup A \backslash B, \mu)(a) = \rho(A, \mu)(a)$ for all $a \in A \backslash B$.*

All alternatives in a cluster of clones contain the same information as comparators for evaluation of the out-of-cluster elements $a \in A \backslash B$. Hence the exclusion or inclusion of some of them should not affect the index-weights of the out-of cluster elements. Note that this also implies that the total weight of all elements in the cluster remains unchanged: $\rho(B' \cup A \backslash B, \mu)(B') = \rho(A, \mu)(B)$.

A set $B \subseteq A$ is a **cluster of equal clones** at $\mu$ if it is a cluster of clones and, for all $b, b' \in B$, $M(b, b') = 0$.

**Proposition 14** *1. If $\rho$ satisfies Null Consistency, then, for any $(A, \mu) \in \mathfrak{D}$ and any $b \in A$ such that $\rho(A, \mu)(b) = 0$ and $B_\rho(A, \mu) \backslash b \neq \varnothing$, then*

$$B_\rho(A \backslash b, \mu) = B_\rho(A, \mu) \backslash \{b\}.$$

2. *If $\rho$ satisfies Null Consistency and Pareto Irrelevance, then $B_\rho$ satisfies **Pareto Independence**. That is, for any $(A,\mu) \in \mathfrak{D}$ and any $b \in A$ such that $b$ is Pareto-dominated in $A$,*

$$B_\rho (A \backslash b, \mu) = B_\rho (A, \mu) \backslash \{b\}.$$

3. *If $\rho$ satisfies Cloning Invariance, then $B_\rho$ satisfies **Independence of Equal Clones**. That is, for any $(A,\mu) \in \mathfrak{D}$ and any $\varnothing \neq B' \subseteq B \subseteq A$ such that $B$ is a cluster of equal clones at $\mu$,*

$$B_\rho (B' \cup A \backslash B, \mu) = B_\rho (A, \mu) \cap (B' \cup A \backslash B).$$

A few remarks are in order.

- Note that, in (1), the clause "$B_\rho (A,\mu) \backslash b \neq \varnothing$" cannot be dispensed with, since it is quite possible that $B_\rho (A,\mu) = \{b\}$ even though $\rho (A,\mu)(b) = 0$.
- This issue does not arise in (2), since if $B_\rho (A,\mu) = \{b\}$, $b$ cannot be Pareto dominated.
- In (3), the non-standard restriction to equal clones is necessary; we will show by example in the next section that the standard axiom of Independence of Clones (which drops this clause) will be violated by $B_\rho$ whenever $\rho$ is continuous and dispersed.[8]

# 6 Systematic Departures from Condorcet on the Single-Peaked Domain

Consider profiles supported on the single-peaked domain on three alternatives. These can be desribed as four-tuples $(\alpha, \beta_1, \beta_2, \gamma)$, where

$$\mu(abc) = \alpha,\ \mu(bac) = \beta_1,\ \mu(bca) = \beta_2,\ \mu(cba) = \gamma,$$

summing up to one.[9]

We note first that if neither of the two "extreme" alternatives $a$ or $c$ is favored by a majority of individuals (i.e. if $\alpha, \beta > 0$), which is exactly the case if $b$ is the unique Condorcet winner, $b$ is in fact the unique ordinally admissible alternative. But if the unique Condorcet winner is one of the extremes, it is generally not the unique ordinally admissible alternative, and the "center" alternative $b$ may – indeed: will! - sometimes be chosen. We develop the analysis in a sequence of steps.

[8] For most choice correspondences in the literature to date, the two versions of "Independence of clones" are equivalent. But there are exceptions. The minmax rule, for example, satisfies Independence of Equal Clones but does not satisfy Independence of Clones.

[9] In this notation, alternatives are listed top to bottom; so *abc* is short for $aPbPc$, etc. .

**Claim 15** *If* $\alpha = \frac{1}{2}, \beta_2 > 0, \gamma > 0$ *and* $\rho$ *satisfies Dispersion, then* $B_\rho(\{a,b,c\},\mu) = \{b\}$*.*

Proof. Under the assumptions on the profile,

$$\begin{aligned}\overline{M}(a,b) &= \overline{M}(b,a) = \frac{1}{2};\\ \overline{M}(a,c) &= \alpha + \beta_1;\\ \overline{M}(b,c) &= 1-\gamma.\end{aligned}$$

Since $\beta_2 > 0$, $M(a,c) < M(b,c)$. So the indirect comparison between $a$ and $b$ is in $b's$ favor, while the direct comparison is tied. By Dispersion, $\rho(c) > 0$, hence the indirect comparison carries strictly positive weight. Hence $a \notin B_\rho(\{a,b,c\},\mu)$. Further, $c \notin B_\rho(\{a,b,c\},\mu)$, since $c$ is a Condorcet Loser. Thus $B_\rho(\{a,b,c\},\mu) = \{b\}$. □

From Claim 15, we immediately get:

**Proposition 16** *If* $\rho$ *satisfies Dispersion and Continuity, then there exists a single-peaked profile* $\mu$ *such that* $\alpha > \frac{1}{2}$ *and* $B_\rho(\{a,b,c\},\mu) = \{b\}$*. In particular,* $B_\rho$ *violates Condorcet Consistency.*

**Proof.** Take a single-peaked profile $\mu'$ such that $\alpha' = \frac{1}{2}, \beta_2' > 0, \gamma' > 0$. By Claim 15, $B_\rho(\{a,b,c\},\mu') = \{b\}$. It is straightforward to verify that continuity of $\rho$ implies upper-hemi-continuity of $B_\rho$. Thus $B_\rho(\{a,b,c\},\mu) = \{b\}$ for all $\mu$ in a neighborhood of $\mu'$. The assertion of the Proposition follows. □

**Proposition 17** *If* $\rho$ *satisfies Dispersion and Continuity, then* $B_\rho$ *violates Independence of Clones.*

**Proof.** Consider special case of Proposition 16 in which $\alpha > \frac{1}{2}, \beta_1 = 0$ and $B_\rho(\{a,b,c\},\mu) = \{b\}$. Then the set $\{b,c\}$ is a set of clones. Thus Clone Independence implies $B_\rho(\{a,b,c\},\mu) \cap \{a,b\} = B_\rho(\{a,b\},\mu)$. But the latter equals $\{a\}$ since $\alpha > \frac{1}{2}$, a contradiction. □

To see why it would not be appropriate for Clone Independence *of choice* to hold in this example, note that the assumptions on $\rho$ are perfectly consistent with the assumption of Clone Independence of the index measure $\rho$, for the latter merely requires that the weight of $b$ in $(\{a,b\},\mu)$ be split among $b$ and $c$ in $(\{a,b,c\},\mu)$, which is perfectly consistent with the weight of $c$ being stricly positive in the latter. Thus the inclusion of $c$ "helps" $b$ in comparison to $a$, since $b$ outshines $a$ in the indirect comparison via $c$.

The above Propositions demonstrating some Borda-like behaviour rely only on the weak regularitly assumptions of Dispersion and Continuity. The Borda rule itself is a special case. If we add the reliablity requirement of Pareto Irrelevance on the weighting function, we get something quite Borda-unlike, namely a necessary violation of Monotonicity, one of the classical "paradoxes of voting".

**Condition 18** ***(Monotonicity)*** *If $\mu = \alpha\mu_0 + (1-\alpha)\delta_P$, and $\mu' = \alpha\mu_0 + (1-\alpha)\delta_{P'}$, where $P'$ differs from $P$ only in $b$ being moved up in the ranking, $C(A,\mu) = \{b\}$ implies $C(A,\mu) = \{b'\}$.*

**Proposition 19** *If $\rho$ satifies Dispersion, Continuity and Pareto Irrelevance, then $B_\rho$ violates Monotonicity.*

**Proof.** Take a single-peaked profile $\mu$ such that $\alpha > \frac{1}{2}, \beta_2 > 0, \gamma > 0$. By Proposition 16, $B_\rho(\{a,b,c\},\mu) = \{b\}$. Now consider the profile $\mu'$ such that $\alpha' = \alpha$, $\beta_1' = \beta_1$, $\beta_2' = \beta_2 + \gamma$, $\gamma^{'} = 0$. The profile $\mu'$ results from $\mu$ by a flip in the preference between $b$ and $c$ among all individuals that had $c$ as their top in profile $\mu$. That is, this subset of individual has preference ranking $cPbPa$ in $\mu$, while it has preference ranking $bPcPa$ in $\mu$; the other individuals rankings remain the same. So Monotonicity requires that $B_\rho(\{a,b,c\},\mu') = \{b\}$. But, by Pareto Irrelevance, $\rho(\{a,b,c\},\mu')(c) = 0$ since $\gamma = 0$. So the choice among $a$ and $b$ is decided by the direct majority comparison, which is in favor of $a$. Thus $B_\rho(\{a,b,c\},\mu') = \{a\}$, a contradiction. □

Violations of Monotonicity are certainly counterintuitive at first sight, and ask for explanation. Here is one. The choice of $b$ at profile $\mu$ is perfectly reasonable, as established above. If the weight assignments at the two profiles was the same – i.e. if $\rho(\{a,b,c\},\mu) = \rho(\{a,b,c\},\mu)$ – monotonicity would be verified. But there are good reasons that these weight assignments should not be the same, instantiated here by the Pareto Irrelevance conditions. Obviously, this change in weights might affect the relative evaluation between $a$ and $b$ – that indeed is the very point of such change, one might say. There seems to be no a priori reason why this second effect might not outweight the first, *ceteris-paribus* effect. We conclude that while the monotonicity axioms appears plausible by drawing on a ceteris-paribus intuition, but is not normatively binding in general as it ignores the possibility of *ceteris-non-paribus* effects.

# 7 The Plurality Index

In this section, we propose a particular index satisfying all the above desiderata, the "plurality index" $\pi$. The **plurality index** $\pi$ is simply given by the distribution of individuals' preference tops: $\pi(A,\mu)(e) := \mu(\{P : top_A(P) = e\})$. We will refer to the associated generalized Borda rule $B_\pi$ as the "plurality-weighted Borda rule", or **pluri-Borda rule** for short.

In the language of probabilistic choice, $\pi(A,\mu)$ is the "random dictatorship" lottery, and aggregate pluri-Borda scores $s(a,\pi(A,\mu))$ have the following intepretation. Pick randomly and independently two individuals, $i$ and $j$; the score $s(a,\pi(A,\mu))$ is the probability that individual $i$ prefers $a$ to individual $j's$ top alternative.

The plurality index is appealing for the simplicity of its definition and computation, its intuitive accessibility and analytical tractability, and its apparent *sense.* We do not claim it to be "the" "ideal" index, and it well may not be (if such ideal index exists at all). At the same time, the plurality index does have a simple and illuminanting axiomatic characterization that rest on just two axioms. The first is the "Irrelevance of (Pareto) inferior alternatives" axiom stated above as a central reliability desideratum for indices. The second is the axiom of "Mixture-Invariance".

**Axiom 20** ***(Mixture Invariance)*** *For all $A \in \mathcal{F}(\mathfrak{A})$, all $\mu, \mu' \in \Delta(\mathcal{L}(A))$, and all rational $\alpha \in [0,1]$, $\rho(A, \alpha\mu + (1-\alpha)\mu') = \alpha\rho(A,\mu) + (1-\alpha)\rho(A,\mu')$.*

Mixture invariance asserts strong a strong form of "population consistency".

Mathematically, it reflects a basic structural fact about indices, namely the fact that they are functions from one probability space into another. Mixture Invariance says that the basic algebraic operation in such spaces, the formation of convex combinations, is preserved under $\rho$. From this perspective, Mixture Invariance resembles the Additivity axiom underlying the standard axiomatization of the Shapley value, which reflects the fact that cooperative values are mappings from one linear space to another.

**Theorem 21** *The plurality index is the unique index satisfying Irrelevance of Pareto-Inferior Alternatives and Mixture Invariance.*

**Proof.** The proof is extremely simple, not altogether unlike the proof of the standard characterization of the Shapley value.

Consider first unanimous profiles $\delta_P$ concentrated on a single preference ordering $P$. All alternatives other than $top(A,P)$ are Pareto dominated. Hence $\rho(A,\delta_P) = \delta_{top(A,P)}$.

Now any profile $\mu \in \Delta(A)$ can be written as a convex combination of unanimous profiles, $\mu = \sum_{P\in\mathcal{L}(A)} \mu_P \delta_P$. By Mixture Invariance, $\rho(A,\mu) = \sum_{P\in\mathcal{L}(A)} \mu_P \rho(A,\delta_P) = \sum_{P\in\mathcal{L}(A)} \mu_P \delta_{top(A,P)} = \pi(A,\mu)$.

The converse is straight forward. □

## 8 Summing Up: An Arrowian Possibility Result

We can summarize the thrust of the discussion of the paper and put the contribution of the Pluri-Borda rule in the context of the broader social choice and voting literature by way of an "Arrowian" possibility result. To formulate it, we need to formulate one key desiratum on choice correspondences, "Resoluteness". Resoluteness requires that that the correrspondence select a unique choice "generically" in the following sense, very loosely: "almost everywhere".

**Axiom 22** ***(Resoluteness)*** *For every agenda $A$, $C(A, \mu)$ is single-valued on an open and dense set of profiles $\mu \in \Delta(A)$.*

**Theorem 23** *There exists a choice correspondence that satisfies Ordinal Efficiency, Pareto Independence, Independence of Equal Clones and Resoluteness. In fact, the Pluri-Borda Rule satisfies these four properties.*

We have note the satisfaction of the first three conditions already. In the Appendix, it is shown that the Pluri-Borda rule is resolute.

To the best of our knowledge, it is the only such correspondence in the existing literature. Here are some well-known rules that come close to satisfying the four requirements of the Theorem.

1. The Single-Transferable Vote (a.k.a Alternative Vote) satisfies all requirements except for Ordinal Efficiency; indeed, it is not even ordinally admissible.

2. The minmax rule as well satisfies all requirement except for Ordinal Efficiency; in contrast to the STV, it is ordinally admissible.

3. The Essential Set defined in the next section satisfies all requirements except for Resoluteness. As to Resoluteness, it is easy to see that the Essential set must contain at laest three alternatives in the absence of a Condorcet winner. Since contains an open set, uniqueness is not a dense property.

## 9 The Essential Set, a Condorcetian GBR

As noted above, if a profile admits a Condorcet winner, such winner is ordinally admissible, indeed ordinally efficient. Is it possible to reconcile the requiremenent the Condorcet winners always be chosen ("Condorcet Consistency") with ordinal efficiency?

In other words, do their exist "well-bevaved" ordinally efficient choice correspondences that select Condorcet winners whenver they exist? Without any additional restrictions, the question is trivial, as one could just select, in the absence of a Condorcet winner, ordinally efficient alternatives in some arbitrary or ad-hoc ways. An example is Black's (1948?) rule which selects the Condorcet winner, if a winner exists, and otherwise selects the Borda winner. So one would one want to impose some additional restrictions, such as selection from the top cycle, that exdend the Condorcetian spririt to in a natural manner. Note that the answer to the revised, somewhat open-ended question is not entirely obvious, as some natural candidates such as the minmax rule fail to be ordinally efficient, while being ordinally admissible.

But we can find in the literature a choice correspondence that fits the bill quite elegantly: the Essential Set. The Essential Set is most easily motivate via the notion of a "maximal lottery". A lottery $p$ is a **maximal lottery** for $(A, \mu)$ if $M(p, q) \geq 0$ for all $q \in \Delta(A)$. Maximal lotteries can be viewed as lotteries that are Condorcet winners in an expected sense.

Since $M$ is skew-symmetric, maximal lotteries correspond to the max-min strategies of a zero sum game. Therefore, by the minmax theorem, they always exists (Kreweras 1963, Fishburn 1984). They need not be unique in general, but are always unique when the number of individuals is odd (Laffond et al. (1997).

The Essential Set at $(A, \mu)$ is defined as the set of alternatives that have positive probability for some maximal lottery:

$$ES(A, \mu) = \{a \in A : \ p_a > 0 \text{ for some } p \in ML(A, \mu)\}.$$

Ordinal efficiency of alternatives in the Essential Set is immediate from its definition. Joint ordinal efficiency of the Essential Set itself follows from the convexity of the set of maximal lotteries, which implies the existence of a maximal lottery with maximal support, i.e. support equal to the essential set.

Thus, by Theorem 7, the Essential Set is contained in a GBR. But one can show that it is in fact a GBR. This follows from the following Lemma adapted from Laslier (2000, p. 278).

**Lemma 24** *For any $(A, \mu) \in \mathfrak{D}$, there exists $p \in \Delta(A)$ such that*
*i) the support of $p$ equals $ES(A, \mu)$, and*
*ii) for all $a \in A$, $a \in ES(A, \mu)$ iff $M(a, p) = 0$, and $a \notin ES(A, \mu)$ iff $M(a, p) < 0$.*

Let the set of $p$ described in the Lemma be denoted by $ML^o(A, \mu) \subseteq ML(A, \mu)$.

We thus obtain the following result.

**Proposition 25** *The Essential Set correspondence $ES$ is a Generalized Borda Rule $ES = B_\rho$, where $\rho$ is any selection from $ML^o$.*

In section 6, we have shown that "regular" indices will produce Condorcet inconsistent choices at some profiles. But the Essential Set is Condorcet consistent. So the indices associated with the Essential Set cannot be regular. Indeed, any selection $\rho \in ML$ violates Dispersion, since, for any Condorcet consistent profile $\mu$ in $A$, $\rho(a) = 1$ if $a$ is the Condorcet winner. Likewise, any selection $\rho \in ML$ violates Continuity, in particular: continuity at the uniform distribution $\overline{\mu}$ on $\mathcal{L}(A)$. To see this, consider profiles of the form $(1-\varepsilon)\overline{\mu} + \varepsilon\delta_P$, where $P$ is any preference ranking in $\mathcal{L}(A)$ and $\varepsilon > 0$. All of these profiles are Condorcet consistent, with $top(P)$ being the unique Condorcet winner. Any $\rho \in ML$ will put $\rho(A, (1-\varepsilon)\overline{\mu} + \varepsilon\delta_P) = \delta_{top(P)}$. By Continuity, $\rho(A, \overline{\mu}) = \delta_{top(P)}$, which leads to contradiction since $P$ is arbitrary.[10]

So while the Essential Set has impeccable reliability properties, including Independence of Clones and Pareto Independence (see Laslier 2000), by relying on indices which are arguably too narrow and not as "informative" as they could be, it does not appear to be fully satisfactory as a criterion of ordinal welfarism.

Setting this discussion aside, in view of the many attractive properties of the Essential Set, one would want to find a place for it in the broader scheme of Generalized Borda Rules. Here is one way to do it, via the underlying ML-indices. The following result points out that these indices are exactly the ones which allow alternatives to be relevant as comparators *only* if they are best choices.

**Axiom 26** *(Choice Congruence)* *For all* $(A, \mu) \in \mathfrak{D}$, $rel(A, \mu) \subseteq B_\rho(A, \mu)$.

**Proposition 27** *An index $\rho$ with associated GBR $B_\rho$ satisfies Choice Congruence if and only if $\rho$ is a maximal lottery index. ($\rho \in ML$).*

Note that in view of Proposition 25, this result doesn't characterize the Essential Set exactly, but only a somewhat larger family of correspondences.

**Proof. If.** Fix $(A, \mu)$. Take $p = \rho(A, \mu) \in ML(A, \mu)$. Since $p$ maximizes $M(q, p)$ over $q \in \Delta(A)$, so does any $a$ in in its support $rel(A, \mu)$, which verifies the inclusion $rel(A, \mu) \subseteq B_\rho(A, \mu)$.

For the converse, take any $(A, \mu)$ and take $p = \rho(A, \mu)$. Now by skew-symmetry of $M$, $M(p, p) = 0$. Thus $M(a, p) = 0$ for all $a \in rel(A, \mu)$. By Choice

[10] One might consider allowing indices to be set- valued, and point to the upper-hemicontinuity of $ML$ as a correspondence. While this reasonable for $ML$ as a stochastic choice correspondence as in Fishburn (1984) and Brandl et al. (2016), it does not seem appealing when $ML$ is a correspondence of indices. The point is that, in view of the perfect symmetry of the uniform distribution, arguably any "reasonable" index should assign uniform weight to all alternatives, while $ML(A, \overline{\mu}) = \Delta(A)$.

Congruence, $M(b,p) \leq 0$ for all $b \in B_{\rho}(A,\mu)$. Thus, the dual (minimum) player garantuees 0 by playing $p$, ie. $p = \rho(A,\mu) \in ML(A,\mu)$.$\square$

# 10 Appendix: Remaing Proofs

## 10.1 Characterization of Borda Rule

**Proof of Proposition 2.**

The if-part follows from the majority margin representation of the Borda rule.

The only-if part follows from the following Lemma:

**Lemma 28** *For any* $m, m', m'', m''' \in \mathbb{N}$ *such that* $m+m' = m''+m''' \leq |A|+1$, $t(m) + t(m') = t(m'') + t(m''')$.

**Proof.** By way of contradiction. Consider a quadruple $m, m', m'', m''' \in \mathbb{N}$ such that $m + m' = m'' + m''' \leq |A| + 1$, but $t(m) + t(m') < t(m'') + t(m''')$.

Let $M := m + m' - 1 \leq |A|$.

Consider a profile $\mu$ concentrated on 3 linear rankings $P, P', P''$ with masses $\mu(P) = \mu(P') = \frac{1}{2} - \varepsilon$ and $\mu(P'') = 2\varepsilon$

such that $(a_1 P ... P a_{|A|})$, $(a_M P'..P' a_1 P' a_{M+1} P' ... P' a_{|A|})$ and $P''$ any ordering such that $a_m P'' a$ for all $a \neq a_m$ and, for all $k \leq M$ and $k' > M$, $a_k P'' a_{k'}$.

By construction, for all $k \leq M$ and $k' > M$, $a_k$ is unanimously preferred to $a_{k'}$.

Also, for any $k, k' \leq M$, $M(a_k, a_{k'}) = 2\varepsilon$ iff $a_k P'' a_{k'}$.

Thus $a_m$ is the unique ordinally admissible alternative (irrespective of $\varepsilon > 0$).

While $a_m$ has the $m-th$ rank in $P$, it has the $m'$-th rank in $P'$. Dto. $a_{m''}$ has the $m''-th$ rank in $P$, and the $m''$-th rank in $P'$; let $\widetilde{m}$ denote its rank in $P''$.

The average score of $a_m$ is

$$\left(\frac{1}{2} - \varepsilon\right) t(m) + \left(\frac{1}{2} - \varepsilon\right) t(m') + 2\varepsilon s(1),$$

while the average score of $a_{m''}$ is

$$\left(\frac{1}{2} - \varepsilon\right) t(m'') + \left(\frac{1}{2} - \varepsilon\right) t(m''') + 2\varepsilon s(\widetilde{m}),$$

which is larger if $\varepsilon$ is sufficiently small since $t(m) + t(m') < t(m'') + t(m''')$ by assumption. □

Now, for any $m' = 1$ and any $m \leq |A|$, consider $m'' = 2$ and $m''' = m - 1$. Then the Lemma yields, for all $m : 2 \leq m \leq |A|$,

$$\begin{aligned} t(1) + t(m) &= t(2) + t(m-1), \text{ i.o.w.} \\ t(1) - t(2) &= t(m-1) - t(m); \end{aligned}$$

hence $C^{t-score}$ is in fact the Borda rule. ■

## 10.2 Ordinal Efficiency vs. Joint Ordinal Efficiency

**Proof of Proposition 5.**

Let $A = \{a_1, ..., a_m\}$, $\varepsilon > 0$ sufficiently small, and $\mu \in \Delta(\mathcal{L}(A))$ such that,

1. $\ell = 1, .., m$, $M(a_\ell, a_{\ell+1}) = \varepsilon$, and
2. for $\ell, \ell' \leq m$ such that $|\ell - \ell'| \geq 2$, $|M(a_\ell, a_{\ell'})| < \varepsilon$.

Addition is $mod_m$; for sufficiently small $\varepsilon$, such profiles profiles exist by the McGarvey's (1953) theorem.

First, we verify that each $a \in A$ is ordinally efficient. Indeed, by construction, $a_\ell$ has the largest majority margin over $a_{\ell+1}$ among all $a \in A$. It thus verifies Proposition 3 for the weight vector $(0, .., 1_{a_{\ell+1}}, ...0)$, establishing ordinal efficiency.

Now suppose that $A$ is jointly ordinally efficient (as a set). By (the argument of) Proposition 3, there exists some weight vector $w \in \Delta(A)$ such that

$$A = \arg\max_{a \in A} M(a, w).$$

By the skew-symmetry of $M$, $M(w, w) = 0$. Hence $\max_{a \in A} M(a, w) = 0$. It follows that $w$ is a maximal lottery in $(A, \mu)$ and $A$ the essential set.

By a well-known result due to Laffond et al. (1997), if the number of individuals is odd ,then the unique maximal lottery has support of odd cardinality. If follows that if $|A|$ is even and at least 4, and if the number of individuals is odd, $A$ cannot be jointly ordinally efficient in $(A, \mu)$. ■

## 10.3 Resoluteness of Pluri-Borda

**Proposition 29** *The Pluri-Borda rule $B_\pi$ is resolute.*

We need to show that, for any agenda $A$, $\{\mu \in \Delta_{\mathbb{Q}}(\mathcal{L}(A)) : |B_\pi(A,\mu)| = 1\}$ is open and dense in $\Delta_{\mathbb{Q}}(\mathcal{L}(A))$.

Openness is straightforward; it remains to verify density.

Consider any aggregation problem $(A,\mu)$.

<u>Case 1. Two tops</u>, ie. $|\{\pi(a) > 0\}| = 2$.
In this case, $B_\pi(A,\mu)$ is multi-valued iff $\pi(a_1) = \pi(a_2) = \frac{1}{2}$ for some $a_1, a_2$, and $B_\pi(A,\mu) = \{a_1, a_2\}$. Let $\mu'$ be any unanimous profile with unanimous top $a_1$. The profiles $\mu_\varepsilon := (1-\varepsilon)\mu + \varepsilon\mu'$ are two-tops profiles, with $a_1$ the absolute majority winner. Evidently $B_\pi(A,\mu_\varepsilon) = \{a_1\}$, verifying density.

<u>Case 2. More than two tops.</u>

Take any $a$ among $B_\pi(A,\mu)$. ($\pi(a)$ might be zero, i.e. $a$ may or may not be a top at $\mu$).

Write $\mu = \sum_{i \in \{1,..n\}} \frac{1}{n}\delta_{R_i}$ and construct $\mu' = \sum_{i \in \{1,..n\}} \frac{1}{n}\delta_{R_i'}$ by requiring the following conditions of $R_i'$

1. For any $b \in \{a\} \cup tops(\mu)$ and $b' \in A \setminus (\{a\} \cup tops(\mu))$,

   $$bP_i'b'.$$

   So all the non-tops other than $a$ are moved to the bottom (in arbitrary manner).

2. For any $b, b' \in tops(\mu) \setminus \{a\}$,

   $$bR_i'b' \text{ iff } bR_i'b'.$$

   So these preferences remain unchanged.

3. For any $b \in A \setminus \{top(R_i), a\}$,

   $$aP_i'b.$$

   So if $a$ is the top of $R_i$, $R_i'$ leaves it there. If $a$ is not the top of $R_i$, $R'$ moves $a$ into second place.

**Claim 30** *Let $a \in B_\pi(A,\mu)$. If $\pi(a) \geq \frac{1}{2}$ or if $\pi(b) < \frac{1}{2}$ for all $b \in B_\pi(A,\mu)$, $\{a\} = B_\pi(A,\mu')$.*

This follows from the following subclaims. We write $s_\pi(a,\mu)$ short for the aggregate pluri-Borda score $\sum_{P\in\mathcal{L}(A)}\mu_P s_P(a,\pi(A,\mu))$.

1. For all $b\in A\backslash(\{a\}\cup tops(\mu))$, $b\notin B_\pi(A,\mu')$.

   For such $b$, by the construction of $\mu'$, their pluri-Borda score $s_\pi(b,\mu)$ equals 0.

2. For all $b\in tops(\mu)\backslash\{a\}$, if $0<\pi(b)<\frac{1}{2}$, then

$$\begin{aligned} s_\pi(a,\mu)-s_\pi(b,\mu) &\geq (1-\pi(b))\frac{\pi(a)+\pi(b)}{2}-\pi(b)\left(\frac{\pi(a)+\pi(b)}{2}\right) \\ &= (1-2\pi(b))\left(\frac{\pi(a)+\pi(b)}{2}\right) \\ &> 0. \end{aligned}$$

   The last inequality is strict since both terms are stricly positive. For the first, this follows from the condition $\pi(b)<\frac{1}{2}$, for the latter from $0<\pi(b)$.

3. For all $b\in tops(\mu)\backslash B_\pi(A,\mu)$,

$$s_\pi(b,\mu')\leq s_\pi(b,\mu)<s_\pi(a,\mu)\leq s_\pi(a,\mu').$$

   The outer inequalities follow from the construction of $\mu'$, the inner strict inequality from $b\notin B_\pi(A,\mu)$. Since $s_\pi(b,\mu')<s_\pi(a,\mu')$, $b\notin B_\pi(A,\mu')$.

So if $\pi(b)\geq\frac{1}{2}$ for some $b\in B_\pi(A,\mu)$, let $a:=b$. Since the profile has at least three tops, the Claim follows from subclaims (1) and (2).

On the other hand, if $\pi(b)<\frac{1}{2}$ for all $b\in B_\pi(A,\mu)$, let $a$ be any element of $B_\pi(A,\mu)$. For any $b$ with $\pi(b)<\frac{1}{2}$, $b\notin B_\pi(A,\mu')$ by subclaims (1) and (2). For any $b$ with $\pi(b)\geq\frac{1}{2}$, by assumption $b\notin B_\pi(A,\mu)$. Hence $b\notin B_\pi(A,\mu')$, establishing the Claim. □

By the Claim, if there are more than two tops, there exists $a\in B_\pi(A,\mu)$ such that $\{a\}=B_\pi(A,\mu')$ for the associated $\mu'$. Since $\mu$ and $\mu'$ have the same distribution of tops $\pi=\pi'$, for any $\mu_\varepsilon:=(1-\varepsilon)\mu+\varepsilon\mu'$, and all $b\in A$

$$s_\pi(b,\mu_\varepsilon)=(1-\varepsilon)s_\pi(b,\mu)+\varepsilon s_\pi(b,\mu'),$$

whence it follows that $\{a\}=B_\pi(A,\mu_\varepsilon)$, again verifying density of the set of single-valued profiles. ■

# 11 References


- ARROW, K. J., AND L. HURWICZ (1972): "An Optimality Criterion of Decision-Making Under Ignorance," in Uncertainty and Expectations in Economics, ed. by C. F. Carter and J. L. Ford. Oxford: Basil Blackwell.
- BLACK, D. (1958): The Theory of Committees and Elections. Cambridge: Cambridge University Press.
- BRANDL, F., BRANDT, F. and H.G. SEEDING (2016): "Consistent Probabilistic Social Choice," Econometrica 84 (5), 1839-1880.
- CONGAR, R., AND V. MERLIN (2012): "A Characterization of the Maximin Rule in the Context of Voting," Theory and Decision, 72 (1), 131–147.
- DUTTA, B., AND J.-F. LASLIER (1999): "Comparison Functions and Choice Correspondences," Social Choice and Welfare, 16 (4), 513–532.
- FINE, B., AND K. FINE (1974): "Social Choice and Individual Ranking I," Review of Economic Studies, 41 (127), 303–323.
- FISHBURN, P. C. (1984): "Probabilistic Social Choice Based on Simple Voting Comparisons," Review of Economic Studies, 51 (4), 683–692.
- FISHER, D. C., AND J. RYAN (1995): "Tournament Games and Positive Tournaments," Journal of Graph Theory, 19 (2), 217–236.
- GUERDIJKOVA, A. AND K. NEHRING (2014): "Weighing Experts, Weighing Sources: The Diversity Value, Mimeo UC Davis
- KREWERAS, G. (1965): "Aggregation of Preference Orderings," in Mathematics and Social Sciences I: Proceedings of the Seminars of Menthon-Saint-Bernard, France (1–27 July 1960) and of Gösing, Austria (3–27 July 1962), 73–79.
- LAFFOND, G., J. LAINÉ, AND J.-F. LASLIER (1996): "Composition-Consistent Tournament Solutions and Social Choice Functions," Social Choice and Welfare, 13 (1), 75–93.
- LAFFOND, G., J.-F.LASLIER , AND M. LE BRETON (1997): "A Theorem on Symmetric Two-Player Zero-Sum Games," Journal of Economic Theory, 72 (2), 426–431.
- LASLIER, J.F. (2000): "Aggregation of Preferences With a Variable Set of Alternatives," Social Choice and Welfare, 17 (2), 269–282.
- MAY, K. (1952): "A Set of Independent, Necessary and Sufficient Conditions for Simple Majority Decisions," Econometrica, 20 (4), 680–684.

- MCGARVEY, D. C. (1953): “A Theorem on the Construction of Voting Paradoxes,” Econometrica, 21 (4), 608–610.
- MOULIN, H. (1986): “Choosing From a Tournament,” Social Choice and Welfare, 3 (4), 271–291.
- PEYRE, R. (2013): “La quête du graal électoral,” Images des Mathématiques, CNRS.
- SAARI, D. (1990b): “The Borda Dictionary,” Social Choice and Welfare, 7 (4), 279–317.
- SMITH, J. H. (1973): “Aggregation of Preferences With Variable Electorate,” Econometrica, 41 (6), 1027–1041.
- TIDEMAN, T. N. (1987): “Independence of Clones as a Criterion for Voting Rules,” Social Choice and Welfare, 4 (3), 185–206.
- YOUNG, H. P. (1974): “An Axiomatization of Borda’s Rule,” Journal of Economic Theory, 9 (1), 43–52.
- YOUNG, H. P. (1975): “Social Choice Scoring Functions,” SIAM Journal on Applied Mathematics, 28 (4), 824–838.
- ZAVIST, T. M., AND T. N. TIDEMAN (1989): “Complete Independence of Clones in the Ranked Pairs Rule,” Social Choice and Welfare, 6 (2), 167–173.